

Early oral language and later reading development in Spanish-speaking English language learners: Evidence from a nine-year longitudinal study

Michael J. Kieffer

Teachers College, Columbia University, 525 W. 120th Street, Box 66, New York, NY 10027, USA

ARTICLE INFO

Article history:

Received 8 June 2011

Received in revised form 17 February 2012

Accepted 23 February 2012

Available online 17 April 2012

Keywords:

At-risk learners

Oral language

Reading development

Longitudinal studies

ABSTRACT

Using nationally-representative, longitudinal data on a cohort of Spanish-speaking English language learners in the U.S., this study investigated the extent to which early oral language proficiency in Spanish and English predicts later levels and rates of growth in English reading. Latent growth models indicated that both Spanish and English proficiency in kindergarten predicted levels of English reading in third through eighth grade, but that only English proficiency was uniquely predictive. English productive vocabulary was found to be a better predictor of later English reading than more complex measures, i.e., listening comprehension and story retell, contrary to findings for native English speakers. Oral language did not predict later growth rates. Findings suggest the need for educational efforts to develop oral language during early childhood for this underserved population. Findings further suggest that such early efforts may be necessary, but insufficient to accelerate ELLs' reading trajectories as they move into adolescence.

© 2012 Elsevier Inc. All rights reserved.

The precursors of successful reading develop long before students begin formal reading instruction (Snow, Burns, & Griffin, 1998). Substantial longitudinal evidence suggests that native English-speaking children who have developed higher levels of oral language proficiency by kindergarten are more successful in learning to read in the primary grades than those who enter school with underdeveloped oral language (e.g., Catts, Adlof, & Weismer, 2006; Dickinson & Tabors, 2001; Senechal & LaFevre, 2002; Storch & Whitehurst, 2002). However, far less is known about these relationships for the growing population of students who come from homes in which a language other than English is primarily spoken, a group known as language minority (LM) learners (August & Shanahan, 2006). In particular, English language learners (ELLs), the subset of this larger population that enters school not yet proficient in English, demonstrate disproportionately low English reading comprehension in the upper elementary and middle school grades (e.g., Kieffer, 2008, 2010; NCES, 2009), yet it is unclear to what extent these reading comprehension difficulties are predicted by ELLs' earlier oral language development.

In the most recent comprehensive review of research on the early predictors of later reading proficiency, the National Early Literacy Panel (NELP) (2008) found evidence for 11 early or precursor literacy skills, measured between birth and kindergarten, that had moderate to strong correlations with later literacy abilities. Among these, oral language, defined as the “ability to produce or comprehend spoken language, including vocabulary and grammar,” was classified as one of the five “potentially important variables” (p. viii). These five

variables demonstrated moderate relationships with later reading that were weaker than those of the best predictors, which included alphabet knowledge, phonological awareness and memory, rapid automatic naming, and writing. In responding to the report, some researchers have suggested that the finding of only a “potentially important” role for oral language may not apply to ELLs, for whom early oral language development—in both their primary and second languages—could be particularly essential (Gutierrez, Zepeda, & Castro, 2010; Orellana & D'warte, 2010). Leaders of the panel have responded that the relationships found may indeed depend on ELL status, but that the extant evidence does not make it possible to determine whether this is the case (Lonigan & Shanahan, 2010; see also Schnatschneider & Lonigan, 2010).

The present study was designed to extend the current research base on early predictors of later reading development by investigating the role of early oral language in later reading growth among Spanish-speaking ELLs in the U.S., using nine years of longitudinal data on a nationally representative sample of this population. Specifically, this research focused on the subpopulation of Spanish-speaking language minority learners who were U.S.-educated (i.e., born in the U.S. or immigrated before kindergarten) and who entered U.S. kindergarten classrooms with limited English proficiency (i.e., *initial* English language learners at school entry, whether or not they later were redesignated to fluent English proficient). For this population, the roles of Spanish and English oral language in kindergarten in predicting students' later levels and/or rates of growth in English reading between third and eighth grade were examined. This study further investigated whether vocabulary measures or more complex measures of oral language are better predictors of later reading in this population.

E-mail address: mk3157@columbia.edu.

Early oral language and later reading outcomes in native English speakers

In their meta-analytic review, the NELP (2008) found that oral language proficiency measured in kindergarten had a moderate relationship with later decoding across 50 studies as well as a moderate relationship with later reading comprehension across 23 studies. Interestingly, the NELP (2008) found large differences in the predictive relationship as a function of the oral language measure examined, with overall composite measures (i.e., those that combine assessments of vocabulary, syntax, and listening comprehension in the same test) demonstrating much stronger relationships to later decoding and reading comprehension than measures of individual skills. In addition, listening comprehension measures that require students to integrate vocabulary and grammatical knowledge had a notably stronger relationship to later reading comprehension, compared to measures of receptive vocabulary, which were among the weakest predictors in the oral language domain.

While the vast majority of extant studies measured reading outcomes in first or second grade, several studies suggest that the predictive power of early oral language on reading comprehension persists into later grades (e.g., Catts et al., 2006; Cunningham & Stanovich, 1997; Senechal & LaFevre, 2002; Snow, Porche, Tabors, & Harris, 2007; Storch & Whitehurst, 2002). For instance, Snow et al. (2007) found that kindergarten receptive vocabulary had a strong and stable correlation with reading comprehension scores in Grades 4, 7, and 10 for a sample of monolingual children from low-income backgrounds. Although later reading comprehension was also predicted by more complex language measures, including a researcher-created measure of formal definitional skill and a researcher-created measure of narrative production, these measures had weaker correlations than did vocabulary and they declined in predictive power over time. Similarly, Cunningham and Stanovich (1997) found that first-grade receptive vocabulary exhibited a moderate relationship with reading comprehension measured 10 years later. Despite the accumulation of evidence for the importance of early oral language proficiency to first-language reading development through middle school, far less is known about the role of early oral skills for the long-term reading development of ELLs.

Language and reading development in Spanish-speaking ELLs

The rapidly growing population of ELLs in the U.S., more than 70% of whom come from Spanish-speaking homes (Capps et al., 2005), provides new challenges and opportunities to educators who have previously relied on models of first-language reading development. In particular, the disproportionate prevalence of English reading comprehension difficulties among ELLs in the middle grades (e.g., Kieffer, 2010; NCES, 2009) raises the question of whether the precursors to these difficulties can be identified much earlier. With only a few longitudinal studies to date that have followed ELLs beyond fourth grade (e.g., Mancilla-Martinez & Lesaux, 2010; Nakamoto, Lindsey, & Manis, 2007; Reese, Garnier, Gallimore, & Goldenberg, 2000), this question remains largely open.

Although it is likely that the conclusions of the NELP (2008) will largely hold for ELLs, there are some reasons to suspect that early oral language proficiency may play a somewhat different role in the reading development of Spanish-speaking ELLs than it plays in monolinguals' reading development. First and foremost, ELLs' oral language skills are distributed across two languages, so measures of English oral language proficiency alone may not capture the full range of linguistic resources available to these students in the process of learning to read English (e.g., Genesee, Lindholm-Leary, Saunders, & Christian, 2006; Gutierrez et al., 2010). Indeed, evidence from language-of-instruction studies indicates that improving first-language literacy skills supports second-language literacy outcomes (for reviews, see

e.g., Francis, Lesaux, & August, 2006; Genesee et al., 2006; Slavin & Cheung, 2005). However, the evidence to date on the importance of first-language oral proficiency, as opposed to first-language literacy, for second-language reading is more limited and equivocal; Geva and Genesee's (2006) review, as well as two recent studies (Gottardo & Mueller, 2009; Nakamoto, Lindsey, & Manis, 2008), found more evidence for within-language effects than cross-language effects. As a result, researchers have been skeptical that first-language oral proficiency will predict second-language reading or compensate for underdeveloped oral proficiency in the second language (e.g., Bialystok, 2002; Verhoeven, 1994).

A second reason to suspect that early oral language may play a unique role for ELLs is the great heterogeneity in basic language proficiency found among these learners. Whereas native English speakers without clinical language impairments can be assumed to have acquired a command of commonly used vocabulary and basic grammar by kindergarten, ELLs may vary substantially in their command of such language (e.g., August, Carlo, Dressler, & Snow, 2005). Consider, for instance, the words that Beck, McKeown, and Kucan (2002) label Tier 1 words, or "mostly basic words—clock, baby, happy—rarely requiring instruction in school" (p. 16); while most native English speakers and many ELLs arrive at school knowing these words, other ELLs do not. The substantial limitations in early oral language in English demonstrated by many ELLs (e.g., Manis, Lindsey, & Bailey, 2004; Swanson, Rosston, Gerber, & Solari, 2008) could seriously constrain their English reading development. At the same time, however, ELLs, as a group defined by limited English proficiency, do not include those who have attained English oral language skills in the upper range of the distribution, so it is also possible that this constrained variation could lead to weaker correlations between English oral language and later reading in this population than in native English speakers.

Third, taking a more ecological view of reading development, ELL status is frequently confounded with low socioeconomic status (Capps et al., 2005; Cosentino de Cohen, Deterding, & Chu Clewell, 2005) as well as limited access to educational resources supporting reading development (Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003), particularly for Spanish speakers. As a result of this multiplicity of risk factors, Spanish-speaking ELLs may be more vulnerable to the negative effects of lower levels of language skills than their more advantaged counterparts. Although the effects of ELL status, first language background, and SES cannot be easily disentangled, there is a need—at minimum—to include statistical controls for SES when estimating the relationship between early oral language and later reading.

In addition to the possibility that the role of oral language development, in general, differs by language background, it is also conceivable that the subcomponents of English oral language proficiency play differentially important roles for Spanish-speaking ELLs and native English speakers. In particular, vocabulary measures that tap more decontextualized language knowledge may have greater specificity in identifying important English language weaknesses in this population, as compared to global listening comprehension measures that offer more contextual support. This would converge with a growing body of research with older ELLs that highlights limited English vocabulary knowledge as a very common source of difficulties with English reading comprehension (e.g., August et al., 2005; Garcia, 1991; Hutchinson, Whiteley, Smith, & Connors, 2003; Lesaux, Crosson, Kieffer, & Pierce, 2010; Lesaux & Kieffer, 2010), including a few longitudinal studies (Mancilla-Martinez & Lesaux, 2010; Nakamoto et al., 2007). In addition, whereas grammatical knowledge appears to be strongly predictive of later reading for native English speakers (National Early Literacy Panel, 2008), a few studies have demonstrated a weaker relationship between grammatical knowledge and reading for ELLs in the elementary grades (e.g., Jongejan, Verhoeven, & Siegel, 2007; Lipka & Siegel, 2007).

Current study

The current study investigated the role of early oral language in predicting later reading comprehension for Spanish-speaking ELLs in the U.S. To address this aim, data on the language and reading development of a nationally representative sample of Spanish speakers who entered school as ELLs and were followed from kindergarten through eighth grade was used. Specifically, the study examines the relationship of Spanish and English measures of vocabulary, listening comprehension, and story retell (as well as a composite of these three measures) in kindergarten with levels and rates of growth in English reading comprehension between grades three and eight. By providing insight into these relationships, this study aims to build the research base on second-language reading development that can ultimately inform efforts to improve early childhood education for linguistically diverse populations.

Based on findings from prior research, it was hypothesized that English oral language measures in kindergarten would demonstrate moderate relationships with later levels of English reading comprehension as well as moderate relationships with later rates of growth in reading comprehension. It was further hypothesized that when included together in a single model, that Spanish oral language measures would not make a unique contribution to English reading comprehension, beyond the contribution of English oral language. Finally, it was hypothesized that among oral language measures, that vocabulary measures would be more predictive of later reading comprehension than listening comprehension and story retell measures.

The following specific research questions guided this study:

- 1) Do kindergarten levels of oral language development in English and/or Spanish predict later levels and/or rates of growth in English reading for Spanish speakers who enter school as English language learners? Does Spanish oral language development uniquely predict later levels and/or rates of growth in English reading, after controlling for the effects of English oral language development?
- 2) Among individual indicators of early oral language development in English and Spanish, is productive vocabulary more predictive of later levels and/or rates of growth in English reading compared to more complex measures (i.e., listening comprehension, or story retell measures) in this population?

Method

Dataset

This study uses data from the *Early Childhood Longitudinal Study—Kindergarten* cohort (ECLS-K; Tourangeau, Lê, Nord, & Sorongon, 2009), a study conducted by the National Center for Education Statistics (NCES) that examines U.S. school children's academic and behavioral development between kindergarten and eighth grade. ECLS-K was designed to capture information on a wide range of child, home, and school characteristics through student assessment, parent interviews, teacher interviews, and principal surveys. In conducting ECLS-K, NCES used a multi-stage probability sampling design to select a cohort of students that was nationally representative of students entering kindergarten in the 1998–99 school year. The cohort of students participated in data collection on six occasions over the nine years of the study: fall of kindergarten, spring of kindergarten, spring of first grade, spring of third grade, spring of fifth grade, and spring of eighth grade. ECLS-K is one of the first national longitudinal studies to incorporate ELLs fully and appropriately into cognitive assessment and other aspects of data collection and, as such, is well suited to addressing these research questions.

Analytic sample

To address the research questions about the population of U.S.-educated, Spanish-speaking ELLs, the current study was conducted with a sub-sample of 295 students participating in ECLS-K who met the following criteria: 1) parental report of Spanish as the primary home language, 2) limited English proficiency in the fall of kindergarten, i.e., scoring below the proficiency composite cut-score on an English oral language proficiency measure (see details below), and 3) had a valid sampling weight indicating that they had available data on English reading achievement for Grade 3, 5, and 8.¹ Together, the first two criteria also ensured the availability of oral language development scores in students' L1; ECLS-K did not assess students from other language backgrounds in their L1, nor did ECLS-K assess the Spanish skills of Spanish-speaking students who scored above the cut-score on the English oral language measure. The third criteria ensured that the relationship between early oral language and later reading achievement through Grade 8 could be investigated. Although this later criteria excluded students who dropped out of the study over time, ECLS-K provides longitudinal sampling weights that, when included appropriately in analyses, can account for attrition and non-response over time, as well as for the purposeful over-sampling of specific groups (e.g., students attending private schools). All analyses reported below included the appropriate longitudinal sampling weights and thus can support generalizations to the population of Spanish-speaking students who entered kindergarten with limited English proficiency in 1998.

It is worth noting that this definition of Spanish speakers who entered school as ELLs is time-invariant and thus different from the definitions often used by schools and districts, in which “ELL” or “limited English proficient” is a temporary label that children lose when they gain English proficiency. In this way, the population of interest may be more precisely, if less succinctly, referred to as “language minority learners with initially limited English proficiency” (Kieffer, 2008). This longitudinal definition has the benefit of avoiding the major cohort effects that can occur when only the subset of students with an “ELL” designation at a particular later point in time is included. It is also worth noting that the dataset does not include students who immigrated to the U.S. after kindergarten; for this reason, the population of Spanish-speaking ELLs studied is further specified as “U.S.-educated ELLs,” not to be confused with newcomer ELL populations who have been educated outside the U.S. to some extent prior to entering U.S. schools.

The resulting analytic sample was balanced for sex (148 boys and 147 girls). On average, the students came from low socioeconomic backgrounds; they scored 0.78 standard deviations below the national mean on the SES composite index (see below) in kindergarten, with similar SES levels in Grades 1, 3, 5, and 8. On average, participating students were 5 years, 7 months old in the fall of kindergarten ($SD = 4$ months); 9 years, 2 months old in Grade 3 ($SD = 4$ months); 11 years, 1 month old in Grade 5 ($SD = 5$ months); and 14 years, 2 months old in Grade 8 ($SD = 5$ months).

Measures

English reading achievement

Children's overall reading achievement in English was assessed at each testing occasion using a two-stage adaptive reading test assembled by a panel of content experts, with items from published standardized tests, from the National Assessment of Educational Progress (NAEP), and from two earlier longitudinal studies conducted by NCES, the National Education Longitudinal Study of 1988 and the Education Longitudinal Study of 2002 (Najarian, Pollack, & Sorongon, 2009).

¹ In addition to the students removed for missing these scores, 23 students were excluded for missing data on the oral language proficiency measures in kindergarten.

Drawing on and extending the NAEP framework, the test assessed six types of reading comprehension skills (initial understanding, developing interpretation, personal reflection and response, developing a critical stance, and evaluating complex syntax) as well as vocabulary and, in earlier grades, basic reading skills (familiarity with print, recognition of letters, recognition of phonemes, and decoding). A greater proportion of difficult items targeting reading comprehension was included at later grade levels than at earlier grade levels to avoid ceiling effects and to capture age-appropriate skills, and Item-Response Theory (IRT) methods were used by NCEES to create vertically-linked scaled scores that would be comparable over time. Najarian et al. report high estimates of IRT-based reliability (ranging from .87 to .96, by wave) and validity evidence, including high correlations with other standardized measures of reading and lower correlations with mathematics and science measures during field testing. The IRT-derived theta scores were used, because they are vertically linked to be comparable across rounds of data collection. Najarian et al. report evidence supporting the inference that these theta scores measure a unidimensional construct consistently over time, including evidence that the common linking items function the same over time, strong IRT-based fit statistics for the items in the reading measure, and appropriate correlations across waves of data collection. The current study focused on reading achievement in the spring of Grade 3, Grade 5, and Grade 8. It is worth noting that ECLS-K also assessed students in reading during the kindergarten and Grade 1 testing occasions; however, ELLs were systematically excluded from taking the reading assessment at one, two, or all of these occasions, because their oral English proficiency was determined to be too limited for appropriate inclusion in the English reading assessment.

English oral language

Students' oral language development in English was assessed in the fall of kindergarten with three of the six subtests from the English Pre-Language Assessment Scales 2000, Form C (English PreLAS; Duncan & DeAvila, 1998). These subtests were selected in consultation with a panel of experts and outside consultants convened by ECLS-K staff; criteria for selecting the measure included "known psychometric properties including predictive validity and face validity among experts in the field" as well as "widespread use and acceptance for the age group" and similarity in format to the cognitive assessments used in ECLS-K (U.S. Department of Education, 2002, p. 2–21, 2–22).

The first subtest was a *listening comprehension* task referred to as "Simon Says," in which students were asked to respond to simple directives (e.g., *touch your ear, point to the middle of the paper, put one hand on top of the other*) in English by taking an action. The directives were designed to include those likely to be encountered in kindergarten classrooms and to include vocabulary that was limited to parts of the body and items commonly encountered in household and pre-school environments (e.g., *pencil, floor, paper, door*). This task consisted of 10 items, each scored at one point per item.

The second subtest was a *productive vocabulary* task referred to as "Art Show," in which students were asked to provide concrete nouns and basic verbs without inflectional markers to describe pictures. The first portion of the test asks students to name a picture, while the second portion of the test asks students to name a picture and then tell its function or purpose (e.g., *What's this? A book. What can you do with it? Read it.*) This task also consisted of 10 items, each scored at one point per item.

The third subtest was a *story retell* task, referred to as "Let's Tell Stories," in which test administrators read a 75-word story aloud, provided 4 pictures representing events in the story, and asked students to provide an oral retelling of the story, which was then scored on its coherence and elaboration as well as the accuracy and complexity of grammar used. This task consisted of two stories, each of which was

scored on a scale from 0 to 5, which was then weighted by 4 to yield a possible range of 0 to 40 points for the subtest.

The composite score was created by summing the scores from each of the three subtests and thus possible values ranged from 0 to 60.² The ECLS-K psychometric report indicates that one of the authors of the test, Dr. DeAvila recommended that a cut-score of 37 or above indicate that students' oral English proficiency was sufficient to complete the ECLS-K cognitive assessments in English (U.S. Department of Education, 2002). Because this cut-score was used to define the analytic sample in the current study as initially limited English proficient, the current sample's scores range from 0 to 36. The ECLS-K psychometric report for the kindergarten through first grade (U.S. Department of Education, 2002) describes good reliability in the current sample, including a split-half reliability estimate of .97 for the composite score as well as a high Cronbach's alpha for the first two subtests (i.e., estimates reported to be "mostly mid .80s to mid .90s" and "very high for subtests with only ten items each"; p. 7–3). These are consistent with the reliabilities reported by the technical guide for this form of the published PreLAS (alpha = .88 and alternate-forms reliability = .89 for "Simon Says"; alpha = .90, alternate-forms reliability = .94 for "Art Show"; Duncan & DeAvila, (2000)). Although reliability estimates were not reported for the story retell subtest for the current sample, the technical guide for the PreLAS reports alternate-forms reliability of .76 and inter-rater reliability of .88.

Spanish oral language

Students' oral language development in Spanish was assessed with the parallel subtests of the Spanish version of the PreLAS 2000, Form C: *listening comprehension* (i.e., "Tio Simon"), *productive vocabulary* (i.e., "La Casita"), and *story retell* (i.e., "Contando Historias"). These subtests are parallel in format and targeted constructs to those of the English PreLAS, but utilized different stories and stimulus pictures. The test developers reported that the Spanish PreLAS was designed for native Spanish-speaking children, in parallel with the English PreLAS with identical "theoretical distinctions and psychometric foundations" (DeAvila & Duncan, 2000; p. 1). ECLS-K reports a split-half reliability estimate of .92 for the composite score and "high internal consistency" in the listening comprehension and productive vocabulary subtests (U.S. Department of Education, 2002, p. 7–4). The technical manual for the published PreLAS reports adequate reliability for the Spanish subtests (alpha = .74 for "Simon Says;" alpha = .81 for "Art Show;" correlations among stories in the story retell task = .87–.90).

Socioeconomic status

Students' socio-economic status (SES) was measured in kindergarten, Grade 1, Grade 3, Grade 5, and Grade 8 using an ECLS-provided continuous variable that draws on several survey items regarding parents' education, occupation, and household income. Because these five time-varying measures of SES were all highly correlated with one another ($r = .66-.93$), they were combined using confirmatory factor analysis (CFA) into a single latent factor representing average SES over time, which was integrated into subsequent analyses. CFA analyses indicated that this approach is reasonable, with factor loadings all greater than .71 and statistically significant. This approach has the additional advantage of avoiding missing data for SES in subsequent analyses; fitting the CFA models using full-information maximum likelihood estimation accounted for missing data at any given occasion and yielded a

² It is worth noting that using this simple linear composite of raw scores has both advantages and disadvantages. On one hand, this is more analogous to the composite measures used by the studies reviewed in the NELP report. On the other hand, forming a latent composite from the observed subscores might have had greater reliability. It also weights the story retell measure more heavily. However, attempts to form latent composites using confirmatory factor analysis led to convergence problems, likely due to the relatively low correlations among the Spanish subtests. Thus, the simple linear composite of raw scores was used instead.

factor score for every student who had a reported SES value for at least one time point, which ultimately included all students in the analytic sample. Nonetheless, it is worth noting that the substantive results reported below were robust to whether SES was included as a latent factor or as a set of five time-varying indicators.

Data analyses

Latent growth modeling was used to investigate the relationship between early oral language in English and Spanish and later status and growth in English reading. Latent growth modeling of longitudinal data has several advantages over other methods for analyzing change over time that are relevant to the present study (Bollen & Curran, 2006). First, by modeling between-student variability in true intercepts directly, partialing out occasion-specific measurement error, these methods can provide more precise estimates of the relationship between predictor variables (i.e., English and Spanish oral language in the current study) and students' true initial status in the outcome (i.e., third-grade levels of English reading) than offered by observed longitudinal correlations. Second, by modeling between-student variability in true slopes, these methods can allow researchers to determine if predictors are associated with variation in students' true rates of growth in the outcome. Third, by modeling growth in a structural equation modeling framework (as opposed to a multilevel or hierarchical linear modeling framework), latent growth modeling allows for the inclusion of other latent factors as covariates, an advantage that was used in the present study to provide a more powerful approach to controlling for socioeconomic status using the five available measurements of SES. Because there was some evidence of non-normality in the oral language proficiency variables, a sandwich estimator (i.e., the Maximum Likelihood Robust option in Mplus 4.0) was used to estimate standard errors that are robust to deviations from normality.

Fig. 1 presents a path model representing the hypothesized latent growth model to address the first research question. The right hand side of the figure displays a three-indicator two-factor measurement model that represents the development of English reading as a linear function of time. The model links hypothesized latent constructs

(displayed as circles) representing true initial (Grade 3) status (λ_0) and true rate of change (λ_1) in English reading to the observed indicators of English reading on each of the three occasions of measurement (labeled Spring Grade 3 Rdg, Spring Grade 5 Rdg, Spring Grade 8 Rdg). In fitting unconditional growth models, the variance in true rate of change was found to be nonsignificant, so this was fixed to 0. Correspondingly, the covariance between true rate of change and true initial status was also fixed to 0, as shown by the absence of a path between these factors in Fig. 1. The bottom left side of the figure presents the five-indicator one-factor measurement model for the SES composite, linking a hypothesized latent construct for students' SES to the observed indicators of SES on each of five occasions of measurement (kindergarten SES through Grade 8 SES). The effect of this covariate on the growth parameters for English reading is taken into account with the inclusion of structural regression paths (represented as single-headed arrows labeled γ_{01} and γ_{11}). The question predictors of interest, observed composites for English and Spanish oral language, are represented as rectangles in the top left of the figure. The first research question was addressed by evaluating the significance of the structural regression paths between these predictors and the latent growth factors for English reading (represented as single-headed arrows labeled γ_{02} , γ_{12} , γ_{03} , and γ_{13}). Residual covariances among the predictors are also included. To address the second research question, a similar model was fitted, with the three subtests for English oral language and three subtests for Spanish oral language replacing the oral language composites.

Results

Preliminary analyses

Prior to fitting the latent growth curve models, preliminary analyses were conducted to describe the observed levels and relationships among the indicators over time. Table 1 presents the correlations, means, and standard deviations for the observed measures of early oral language and later reading achievement, estimated using the appropriate longitudinal sampling weights. As shown, students' kindergarten performance on the composite measures of oral language in

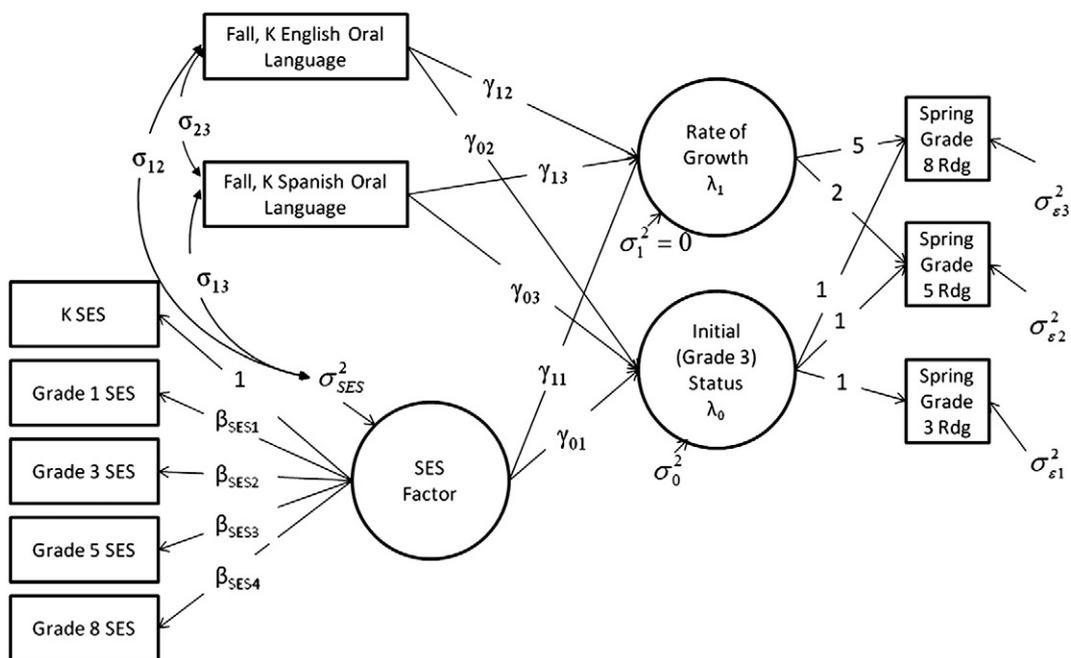


Fig. 1. Path diagram for hypothesized latent growth model for English reading predicted by early English and Spanish oral language, controlling for socioeconomic status.

Table 1

Longitudinal correlations among oral language skills in English and Spanish measured in fall of kindergarten and English reading comprehension measured in Grade 3, 5, and 8, estimated using longitudinal sampling weights ($N = 295$).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. K English oral language composite	–										
2. K English listening comprehension	.87	–									
3. K English productive vocabulary	.81	.71	–								
4. K English story retell	.94	.67	.62	–							
5. K Spanish oral language	.72	.53	.45	.78	–						
6. K Spanish listening comprehension	–.12	.01	–.14	–.15	.15	–					
7. K Spanish productive vocabulary	–.29	–.24	–.20	–.29	.13	.23	–				
8. K Spanish story retell	.83	.60	.53	.90	.91	–.09	–.27	–			
9. Grade 3 English reading	.27	.31	.33	.17	.10	.02	.03	.08	–		
10. Grade 5 English reading	.24	.27	.31	.15	.13	.08	.10	.08	.79	–	
11. Grade 8 English reading	.12	.19	.23	.03	–.02	.02	.03	–.03	.60	.74	–
Means	12.66	3.74	2.98	5.95	22.14	9.77	7.81	4.56	0.50	0.82	1.02
SD	11.95	3.68	2.75	7.00	5.72	0.92	2.12	5.84	0.30	0.24	0.33

Note. K = Kindergarten.

Spanish and English were highly correlated. However, cross-language correlations differed substantially across the subtests of oral language, suggesting the need to investigate these subtests individually.³ Small but non-trivial positive correlations were found between the composite measure of English oral language in kindergarten and later English reading, with a larger magnitude evident for Grade 3 reading than for Grade 5 reading or Grade 8 reading. The English vocabulary subtest demonstrated correlations with later English reading that were slightly higher than the English listening comprehension subtest and notably higher than the English story retell subtest. Correlations between kindergarten Spanish oral language composite and later English reading were quite small for Grade 3 reading and Grade 5 reading and near zero for Grade 8 reading, while the Spanish vocabulary subtest scores had slightly higher correlations with later reading than the other two subtests. As expected, English reading scores were highly correlated over time.

Effects of Spanish and English oral language composites on later English reading

The first research question focused on whether English and/or Spanish oral language in kindergarten, when measured with composite instruments, predicted levels and/or rates of growth in English reading between third and eighth grade, controlling for socioeconomic status. To address this question, a set of latent growth models was fitted to the variance-covariance matrix estimated using the appropriate longitudinal sampling weights. In each model, the predictors of interest were the simple linear composite scores on the Spanish and/or English oral language measures. The fitted models are shown in Table 2.

When included individually, English and Spanish oral language in kindergarten each significantly predicted students' initial (third-grade) status in English reading, as shown in Models 1 and 2 (for English oral language, standardized path coefficient = .25; $z = 3.01$; $p < .01$; for Spanish oral language, standardized path coefficient = .14; $z = 2.00$; $p = .04$). Neither English oral language nor Spanish oral language in kindergarten predicted students' later rates of growth in English reading (for English oral language, $z = -1.40$; $p = .16$; for Spanish oral language, $z = -1.30$; $p = .19$). When English and Spanish oral language are included together in Model 3, English oral language in kindergarten continued to be a significant unique predictor of students' initial

(third-grade) status in English reading (standardized path coefficient = .32; $z = 2.58$; $p = .01$), but Spanish oral language was no longer a significant unique predictor of initial status ($z = -0.95$; $p = .34$). As in Models 1 and 2, neither English nor Spanish oral language predicted students' later rate of growth in English reading when included together (for English oral language, $z = -0.74$; $p = .46$; for Spanish oral language, $z = -0.33$; $p = .74$). Across the models, SES was a statistically significant, positive predictor of initial status, but not of the rate of growth, indicating that students from different SES backgrounds had trajectories that were vertically separated but parallel.

In Model 4, the effect of English oral language on students' initial (third-grade) status in English reading is estimated, after removing the nonsignificant effects of Spanish oral language on initial status and the nonsignificant effects of Spanish and English oral language on rate of growth. When converted into a standardized path coefficient, the effect of English oral language in kindergarten on initial status in reading is .20, a magnitude that indicates a weak relationship according to the conventions of the NELP (2008) and others. However, it is worth noting that this magnitude is roughly the same as the magnitude of the estimated effect of SES (standardized path coefficient = .22).

Given the strong correlation between the English and Spanish oral language composites, there is reason to consider the possibility that a single composite incorporating both Spanish and English oral language skills would be more predictive than either composite individually. Such a composite would serve as a language-general measure that could potentially capture students' communicative resources as they are distributed across two languages better than measures in each language separately. To investigate this possibility, a single latent factor was created to tap the three subtests of the Spanish PreLAS as well as the three subtests of the English PreLAS, and this factor was used to predict initial (third-grade) status and rate of growth in English reading. Although this language-general composite did significantly predict students' initial status in English reading ($z = 2.07$; $p = .04$), this relationship was weaker (standardized path coefficient = .17) than that found for the English oral language composite alone. On closer inspection, this latent composite had appropriately high loadings for only the three English subtests and the Spanish story retell task, so a second language-general composite was created with these four tasks alone; this composite had a similarly weak relationship with students' initial status in reading (standardized path coefficient = .17). Details on these models are available from the author.

Effects of early Spanish and English listening comprehension, vocabulary, and story retell on later English reading

The second research question addressed whether individual measures of early English and Spanish oral language proficiency, as

³ It also worth noting that the Spanish listening comprehension subtest appeared to demonstrate a ceiling effect (i.e., a mean of 9.767 close to the total possible of 10 points and a SD of less than 1 point), presumably because most students had mastered nearly all of the items involving simple directions in their first language. Although the maximum likelihood sandwich estimator used produces standard errors that are robust to this violation of normality, it is worth noting that this ceiling effect may have led to an underestimation of the relationship between Spanish listening comprehension and English reading in subsequent analyses.

Table 2
Selected results for latent growth models for the relationship between initial status and rate of growth in English reading between third and eighth grade, socioeconomic status, kindergarten English oral language composite, and kindergarten Spanish oral language composite ($N = 295$).

		Symbol	Model 1	Model 2	Model 3	Model 4
Fixed effects						
Initial (third grade) status	Intercept	γ_{00}	0.52***	0.46***	0.85*	0.53***
	SES factor	γ_{01}	0.12*	0.14*	0.11*	0.12*
	Fall, K English oral language composite	γ_{02}	0.01**		0.01**	0.01*
	Fall, K Spanish oral language composite	γ_{03}		0.01*	0.00	
Rate of growth	Intercept	γ_{10}	0.11***	0.12***	0.11***	0.10***
	SES factor	γ_{11}	0.01	0.00	0.01	0.00
	Fall, K English oral language composite	γ_{12}	0.00		0.00	
	Fall, K Spanish oral language composite	γ_{13}		0.00	0.00	
Variance components						
Structural model	Variance in initial status	σ_0^2	0.05***	0.05***	0.05***	0.05***
	Correlation between SES factor and fall, K English oral language composite	σ_{12}	.26***	.26***	.26***	.26***
	Correlation between SES factor and fall, K Spanish oral language composite	σ_{13}	.12	.12	.12	.12
	Correlation between fall, K English oral language and fall, K Spanish oral language composite	σ_{23}	.72***	.72***	.72***	.71***
Measurement model	Grade 3 reading	σ_{e1}^2	0.04***	0.04***	0.04***	0.04***
	Grade 5 reading	σ_{e2}^2	0.01*	0.01	0.01*	0.01*
	Grade 8 reading	σ_{e3}^2	0.05***	0.05***	0.05***	0.05***
Goodness of fit						
	–2LogLikelihood		4365.60	4374.22	4363.85	4368.27
	AIC		4429.61	4438.22	4431.86	4430.27
	BIC		4547.59	4556.20	4557.22	4544.57
	χ^2		103.61	105.72	100.36	106.22
	CFI		.91	.91	.91	.91
	RMSEA		.09	.09	.09	.09
	R^2 in initial status		.13	.09	.14	.11
	ΔR^2 for Initial status ^a		.06	.02	.06	.04

Note. Model 1 displays the effect of English oral language alone, model 2 displays the effect of Spanish oral language alone, model 3 displays the simultaneous effects of English and Spanish oral language, and model 4 displays only the uniquely significant effects of English oral language. All Models also included the measurement model for the SES Factor. In all models, variance in rate of growth and covariance between initial status and rate of growth are fixed to zero. Values of 0.00 are due to rounding, rather than fixing effects to zero.

^a These ΔR^2 statistics are in comparison to a model including the effects of SES.

* $p < .05$; ** $p < .10$; *** $p < .001$.

distinguished by three common types of instruments, predicted students' levels and/or rates of growth in later English reading. Specifically, the relative predictive strengths of three subtests of the PreLAS—listening comprehension, productive vocabulary, and story retelling—were investigated in Spanish and English. As with the models that addressed the first research question, a set of latent growth models was fitted, as shown in Table 3: Model 5 included English predictors, Model 6 included Spanish predictors, Model 7 included English and Spanish predictors together, and Model 8 represented the final model in which nonsignificant effects have been removed. As shown in Model 5, among the three measures of early English oral language, productive vocabulary emerged as the only uniquely significant predictor of students' initial (third-grade) status in English reading, controlling for the effects of the other English measures and socioeconomic status. This effect remained significant when the three parallel measures of Spanish oral language were also included, as shown in Model 7. In Model 6, none of the three measures of early Spanish oral language were uniquely significant predictors of initial status or rates of growth, controlling for socioeconomic status. When English and Spanish predictors were included simultaneously, the effect of Spanish story retell on initial status was negative and significant ($z = -2.30$; $p = .02$); however, when the other nonsignificant effects were removed from the model, this effect became nonsignificant ($z = -1.89$; $p = .06$). Thus, in the final model, Model 8, only the significant effect of early English productive vocabulary on students' initial (third-grade) status remained. Across models, none of the English or Spanish subtests predicted students' rates of growth in English reading.

When the significant effect of early English productive vocabulary on later English reading levels in this final model was converted into a standardized path coefficient, it had a magnitude of .29, which approached a moderate size. This indicated a substantially stronger predictive relationship than that found for the English oral language

composite (.20), for the Spanish oral language composite (.14), and for the language-general composite (.17). As a point of comparison, the effect of SES had a magnitude of .20.

Discussion

The current study was designed to investigate the extent to which early English and Spanish oral language predicts later English reading development for the large, growing, and underserved population of Spanish-speaking ELLs in the United States. Building on the relatively robust research base on the moderate role of early oral language in reading development for monolingual children (National Early Literacy Panel, 2008), this research sought to investigate whether similar conclusions held for students learning to read in their second language. Longitudinal data from a nationally representative sample followed for nine years were used to begin to address the widely acknowledged limitations in the empirical research base on ELLs' reading development (e.g., August & Shanahan, 2006; Gutierrez et al., 2010; Lonigan & Shanahan, 2010). The study yielded four major findings, each with theoretical implications for understanding second-language reading development as well as practical implications for designing assessment and instruction for young ELLs that can prevent later reading difficulties. In the following four sections, each finding and its implications for research and practice are discussed.

The importance of early oral language development

The first major finding was that the relationship between early English oral language and later levels of English reading was statistically significant and practically meaningful in magnitude. This finding converges with the handful of past studies that have followed Spanish-speaking ELLs beyond fourth grade (Mancilla-Martinez & Lesaux, 2010; Nakamoto et al., 2007; Reese et al., 2000). The effect

Table 3

Selected results for latent growth models for the relationship between initial status and rate of growth in English reading between third and eighth grade, socioeconomic status, kindergarten English oral language subtests, and kindergarten Spanish oral language subtests ($N = 295$).

		Symbol	Model 5	Model 6	Model 7	Model 8
Fixed effects						
Initial (third grade) status	Intercept	γ_{00}	0.49***	0.36	0.18	0.51***
	SES composite	γ_{01}	0.10	0.15*	0.09	0.11*
	Fall, K English listening comprehension	γ_{04}	0.01		0.01	
	Fall, K English productive vocabulary	γ_{05}	0.02*		0.02*	0.02**
	Fall, K English story retell	γ_{06}	0.00		0.01	
	Fall, K Spanish listening comprehension	γ_{07}		0.01	0.02	
	Fall, K Spanish productive vocabulary	γ_{08}		0.01	0.02	
	Fall, K Spanish story retell	γ_{09}		0.01	−0.01*	
	Rate of Growth	Intercept	γ_{10}	0.11***	0.12*	0.13*
	SES composite	γ_{11}	0.01	0.00	0.01	0.00
	Fall, K English listening comprehension	γ_{14}	0.00		0.00	
	Fall, K English productive vocabulary	γ_{15}	0.00		0.00	
	Fall, K English story retell	γ_{16}	0.00		0.00	
	Fall, K Spanish listening comprehension	γ_{17}		0.00	0.00	
	Fall, K Spanish productive vocabulary	γ_{18}		0.00	0.00	
	Fall, K Spanish story retell	γ_{19}		0.00	0.00	
Variance components						
Structural model	Variance in initial status	σ_0^2	0.05***	0.05***	0.04***	.05***
Measurement model	Grade 3 reading	σ_{E1}^2	0.04***	0.04***	0.04***	.04***
	Grade 5 reading	σ_{E2}^2	0.01*	0.01	0.01*	.01*
	Grade 8 reading	σ_{E3}^2	0.05***	0.05***	0.05***	.05***
Goodness of fit						
	−2LogLikelihood		8415.79	8438.50	8404.00	8450.82
	AIC		8539.79	8562.51	8540.00	8554.82
	BIC		8768.38	8791.10	8790.72	8746.54
	χ^2		142.28	151.27	131.14	163.40
	CFI		.92	.93	.91	.92
	RMSEA		.07	.07	.07	.07
	R^2 in initial status		.18	.10	.23	.13
	ΔR^2 for initial status ^a		.10	.02	.15	.06

Note. Model 5 displays the effect of English oral language subtests alone, model 2 displays the effect of Spanish oral language subtests alone, model 3 displays the simultaneous effects of English and Spanish oral language subtests, and model 4 displays only the uniquely significant effects of English picture vocabulary. All models also included the measurement model for the SES factor as well as structural covariances among the SES Factor, English and Spanish oral language subtests. In all models, variance in rate of growth and covariance between initial status and rate of growth are fixed to 0. Values of 0.00 are due to rounding, rather than fixing effects to zero.

^a These ΔR^2 statistics are in comparison to a model including the effects of SES.

* $p < .05$; ** $p < .10$; *** $p < .001$.

sizes for this relationship (standard path coefficients = .20 to .29), controlling for socioeconomic status, can be considered small to moderate in magnitude. The larger estimate (.29), in particular, is roughly comparable to the average estimate for monolinguals offered by the NELP (2008) ($r = .33$) and thus support the notion that the NELP (2008) conclusion that oral language is a moderate predictor of later reading outcomes can be extended to Spanish-speaking ELLs. As evidence of its practical importance, this effect size was also found to be comparable to or larger than the well-known relationship between SES and reading level. This finding provides valuable support for the importance of early language development in placing ELLs on the pathway to success with reading comprehension through the upper-elementary and middle school grades. At the same time, the similarity in magnitude between these estimates for ELLs and native English speakers suggests that early English oral language development may not be dramatically more important for ELLs than for their peers, as is sometimes suggested.

One practical implication of this finding is that measuring ELLs' oral language in early childhood can provide valuable, if incomplete, information about students' relative risk for developing later reading difficulties. The moderate predictive power of oral language measures suggests that such measures should supplement, rather than replace, measures of other early literacy skills known to predict ELLs' reading, such as phonological awareness, letter knowledge, and working memory (for a review, see Lesaux, Koda, Siegel, & Shanahan, 2006). A second implication is that preschool and kindergarten classrooms must enhance the opportunities for oral language development offered to ELLs as part of regular classroom instruction, in line with the consensus

recommendations offered by the National Academy of Education (2009) among others.

Vocabulary tests may be more predictive than complex measures for ELLs

In a more surprising finding, the English productive vocabulary measure used was a better predictor of later levels of English reading than more complex English language measures, contrary to the conclusion for monolinguals offered by the NELP (2008). When included together with a listening comprehension and a story retell measure—both of which required students to integrate grammar and vocabulary skills—the productive vocabulary measure was the only predictor to demonstrate a significant unique contribution to later levels of English reading. In addition, this single productive vocabulary measure was almost 50% more predictive than a composite of the three measures, contrary to the NELP (2008) findings and to common expectations that a composite will typically be a more reliable predictor than a single sub-test.

Three potential explanations for this divergence in findings are possible, each of which raises questions for future research. First, it is possible that this finding is specific to the subtests of the PreLAS used. Although the PreLAS was widely used in practice and had evidence of solid psychometric properties when selected by a panel of experts during the planning of ECLS-K in the late 1990s, the content and validity of this measure has since been criticized by applied linguistics (e.g., MacSwan & Rolstad, 2006; MacSwan, Rolstad, & Glass, 2002). In addition, the PreLAS does not have the same level of evidence establishing the separability and validity of the individual

vocabulary, listening comprehension, and story recall subtests as do some of the more well-established measures of vocabulary, grammar, and listening comprehension, including many of those used in the studies reviewed by the NELP (2008). Further, this measure was originally designed for the broader purpose of informing decisions about whether ELLs have sufficient English proficiency to benefit from mainstream English instruction, rather than to distinguish specific aspects of oral proficiency that predict their reading development or risk for reading difficulties. As a result, the listening comprehension and story retell tasks may have tapped basic and general communicative competence to a greater extent than many of the clinical and research language measures used in the NELP studies (e.g., the Clinical Evaluation of Language Fundamentals; Wiig, Semel, & Secord, 1992), whereas many of the latter measures tapped more explicit and metalinguistic aspects of grammatical knowledge. The PreLAS composite also weights the story retell measure more heavily than the other subtests, which may not have been warranted. Thus, this finding is subject to replication with different measures. Future research is necessary to investigate not only which constructs best predict later reading development for ELLs, but also which operational approaches to measuring those constructs yield the most predictive power.

A second possible explanation lies in the developmental periods investigated and the relative importance of different skills during those periods. Given that the vast majority of studies reviewed by the NELP (2008) followed students into first or second grade, their conclusion about the relative weak predictive power of early vocabulary may not be applicable to predicting performance in much later grades. It is possible that the relative predictive power of early vocabulary measures is better maintained as students grow older and leave the primary grades, whereas the predictive power of grammar or complex measures declines more with age. Such a changing role for these different aspects of oral language proficiency is consistent with the increasing knowledge demands of the texts students encounter as they move into the upper elementary and middle school grades (e.g., Chall, 1983; Jetton & Alexander, 2004); to the extent that vocabulary is a proxy for students' range of conceptual knowledge about various domains, its predictive power will likely grow over time as knowledge demands increase. This explanation has some support from the few studies that have followed monolingual and Spanish-speaking ELLs beyond the primary grades (e.g., Cunningham & Stanovich, 1997; Mancilla-Martinez & Lesaux, 2010; Reese et al., 2000; Snow et al., 2007), which indeed suggest that early vocabulary continues to have at least a moderate relationship with reading achievement many years later. Although the studies conducted with ELLs did not directly compare early vocabulary with listening comprehension or grammar in predicting post-primary outcomes, Snow et al.'s (2007) study of monolingual children from low-income families did find that a kindergarten receptive vocabulary measure retained the strong magnitude of its correlation with later reading comprehension through Grade 7 and Grade 10 to a greater extent than did a narrative production task or a definitional skill task. That said, the NELP (2008) did not find differences in the existing literature between shorter term and longer term longitudinal studies (Lonigan & Shanahan, 2010), suggesting that future research is needed to settle this question.

A third explanation for the divergence between the current findings and the conclusions from the NELP (2008) is that there are meaningful differences in reading development between Spanish-speaking ELLs and native English speakers. For Spanish-speaking ELLs, complex measures may have less specificity in identifying individual differences that will constrain reading development, if they provide greater contextual support and thus more opportunities to score well on the task, obscuring important deficiencies in English proficiency. A complementary explanation is that the errors in English grammar production that ELLs are likely to demonstrate on a complex proficiency task (such as the story retell task in the current study) may indicate certain real limitations in English proficiency,

but that such limitations have relatively little importance to reading, which involves receptive and arguably less precise knowledge of grammar. Indeed, there is substantial evidence from studies of second-language acquisition in adults of persistent errors in morpho-syntactic features such as inflections and subject-verb agreement (Gass & Selinker, 2008), even among those who are academically successful in their second language; their success, despite flawed English grammar, can be interpreted as evidence that complete mastery of English grammar may be unnecessary for reaching functional levels of reading competence. Such an explanation would also be consistent with the findings from a few studies for the relatively limited predictive power of explicit syntactic awareness tasks for explaining ELLs' reading outcomes in the elementary grades (e.g., Jongejan et al., 2007; Lipka & Siegel, 2007).

Pending replication and confirmation, this finding raises practical questions about the relative emphasis in preschool and primary grade classrooms that should be placed upon vocabulary building activities, compared with activities to promote grammatical development and global listening comprehension. The NELP (2008) interpreted the relatively weaker relationship between simple vocabulary measures and reading outcomes as support for an instructional focus that goes beyond vocabulary alone. They conclude that instructional attention to vocabulary may be important but unlikely to be sufficient for oral language development. While the current findings do not contradict this cautious recommendation or its extension to ELLs, they do imply that there may be a danger in under-emphasizing vocabulary as a precursor literacy skill to be measured and targeted. Given constraints on instructional time, emphasizing the recommendation to "go beyond vocabulary" may inadvertently lead to little time devoted to creating environments and activities that can promote vocabulary learning.

Early Spanish predicts later reading, but not uniquely

The third major finding was that early Spanish oral language demonstrated a significant zero-order correlation with later levels of English reading comprehension, but that this relationship was no longer significant when controlling for early English oral language. When combined with the finding that the early oral language composites for Spanish and English were highly correlated with each other, this suggests that much of the variation in English reading comprehension that is explained by Spanish skills is the same variation explained by English skills. This finding supports a general notion of a common underlying proficiency that is not language-specific (e.g., Cummins, 1979) but is demonstrated on both first- and second-language measures of oral proficiency. This finding also provides some evidence to support the conclusions of Verhoeven (1994), Bialystok (2002), and others in emphasizing the limits of first-language vocabulary and grammar in supporting second-language reading development.

At the same time, this finding should be understood in the context of ELLs' opportunities to learn to read in their first and second languages. It may be that the predictive power of Spanish oral language will be greater for those students who are receiving literacy instruction in their primary language (e.g., Gottardo & Mueller, 2009), unlike the majority of ELLs in the U.S. (e.g., Crawford, 2004). Indeed, the power of oral language skills in predicting later reading is due not only to their longitudinal relationships with later oral language skills which are required for later reading comprehension, but also due to the ways in which well-developed oral language provides students with access to literacy instruction (i.e., allows them to understand what their teacher is saying). For instance, in English-only classrooms, well-developed English oral proficiency may be necessary for first graders to understand their teachers' delivery of systematic phonics instruction in an efficient and deep way, or for third graders to understand instruction in reading comprehension strategies. However, if such instruction is primarily in Spanish or delivered with substantial Spanish support, Spanish oral language proficiency

is much more likely to be important to developing literacy. The relatively limited information on language of instruction reported by ECLS-K makes it difficult to investigate these questions fully with the current dataset, but future research investigating these questions is certainly warranted.

Early language development is not enough for later reading growth

The fourth major finding was that early oral language, whether assessed in English or in Spanish, did not predict later rates of growth in English reading between third and eighth grade. Although students with well-developed early vocabularies had substantially higher reading levels in third grade, compared to their counterparts with under-developed early vocabularies, their trajectories were parallel to their counterparts through eighth grade. That is to say, early vocabulary-related differences in reading levels were maintained through eighth grade, but did not increase. It appears that developing early oral proficiency may provide an advantage in the primary grades, but does not necessarily lead to a pattern of Matthews Effects in which the “rich get richer” and those with weak oral skills fall increasingly further behind (Stanovich, 1986) over the upper-elementary and middle school grades.

This finding contrasts with the findings of Nakamoto et al.'s (2007) study of Spanish-speaking ELLs, which found that an English oral language composite in first grade had a significant, negative association with initial linear rates of growth in English reading comprehension between first and sixth grade as well as a significant, positive association with curvature. Together, their results indicated that ELLs with well-developed early oral language had slower initial growth but less deceleration over time, compared to ELLs with lower levels of early oral language. One explanation for this divergence of findings may be the different developmental periods studied and specifications of growth available. Because the current study did not examine reading growth before the spring of third grade, the negative effect of oral language on early growth may not have been observed. In addition, the inclusion of only three testing occasions for reading in the current study made it impossible to model curvature and thus detect effects on this growth parameter. In addition, this divergence may be due to differences in the samples; the participants in Nakamoto et al. were drawn from transitional bilingual programs within a single school district in a Texas border town, as opposed to the nationally representative sample in the current study. Finally, it is possible that this divergence is due to differences in the oral language or reading measures used. With so few studies to date that have investigated effects of oral language on reading growth in ELLs, there is a clear need for more research to explore this question across a range of developmental periods using various measures.

The findings from the current study support the developmental notion that early oral language development may be necessary for learning to read, but that it is not sufficient to support rapid later growth in reading achievement in the upper-elementary and middle school grades (e.g., Chall, 1983; RAND Reading Study Group, 2002; Snow, Griffin, & Burns, 2005). More rapid growth in these grades than was observed in this study would be necessary for Spanish-speaking ELLs to close achievement gaps with their White, monolingual counterparts (e.g., Kieffer, 2008, 2010; NCES, 2009). As growing research on adolescent literacy indicates (for a review, see Carnegie Council on Adolescent Literacy, 2010), there is a wide range of skills, knowledge, and practices—both within and outside of the domain of oral language—involved in successful reading comprehension of upper-elementary and middle school texts. This finding upholds the general recommendation that concerted efforts to develop oral language should not be the responsibility of early childhood educators alone, but must be sustained and differentiated as students progress through the grades.

Limitations and future research

By drawing on existing data from a large-scale, long-term longitudinal research project, this study offers many advantages over prior research, but also has several limitations. First, as noted above, the findings from the current study may be specific to the individual language and reading measures selected by NCES. Future studies would be well-served by including multiple measures of each oral language skill as well as multiple measures of reading comprehension to facilitate the investigation of relationships among these constructs, while accounting for measurement error, using techniques such as structural equation modeling.

Second, the questions that could be investigated by this study were limited by the collection by NCES of only quantitative assessment and survey data. The description of early oral language would have benefitted from the collection of qualitative observational data on students' actual language use and literacy practices in various contexts, including both home and school. Future large-scale longitudinal studies would benefit from more mixed-methods designs that allow for the triangulation of findings among different data sources and types.

Third, NCES chose to limit the collection of data on English oral proficiency to language minority learners, making it impossible to analyze directly whether the effects of early English oral language skills differ between these learners and native English speakers. In addition, NCES only assessed first-language oral proficiency in Spanish speakers who entered school with initially limited English proficiency. As a result, it was not possible to determine whether different effects for first-language oral proficiency would be found for students who were initially bilingual or for ELLs from other language backgrounds. On a more technical note, it is also possible that this sampling decision led to a restriction of range in the English oral proficiency score that attenuated correlations to some extent. Future comparative research on the full range of language minority learners is needed to shed greater light on these questions.

Finally, the focus on early childhood in the ECLS-K study led NCES to collect only three waves of reading assessment data between Grades 3 and 8, at wider intervals than during kindergarten and Grade 1. As noted above, this made it necessary to specify a linear growth trajectory, when it is likely that growth during this period would be better described as curvilinear (e.g., Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Nakamoto et al., 2007). Although the focus of ECLS-K on early childhood is clearly warranted given its mission and purposes, future large-scale studies that provide a more detailed look at language and reading growth through more frequent data collection during the upper-elementary and middle school years are also needed. The lack of effects of early oral language on students' rates of reading growth between third through eighth grade begs the question of which cognitive, linguistic, social, and instructional factors do predict differences in growth during this period, raising the need for collecting richer data during the middle grades in future longitudinal studies.

Conclusion

Although future research is clearly needed, the current study provides a step toward building an empirical research base about how ELLs' early oral language development supports their later reading development. Findings from this study support the importance of early English oral language development—particularly vocabulary development—to later English reading performance. At the same time, findings suggest the limitations of focusing on oral language alone and on early development alone. Such an emphasis is likely to be a necessary condition, but it is unlikely to be a sufficient condition to promoting more equal long-term reading outcomes for the increasingly linguistically diverse population of U.S. students.

Acknowledgment

This research was supported, in part, by a grant awarded to the author from the American Education Research Association, which receives funds for its AERA Grants Program from the National Science Foundation under NSF Grant No. DRL-0941014. The opinions expressed here and the author's and do not necessarily reflect those of the granting agencies.

References

- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice, 20*(1), 50–57.
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, New Jersey: Lawrence Erlbaum.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life*. New York, NY: Guilford Press.
- Bialystok, E. (2002). Acquisition of literacy in bilingual children: A framework for research. *Language Learning, 52*, 159–199.
- Bollen, K. A., & Curran, P. J. (2006). *Latent growth curve models: A structural equation approach*. Hoboken, NJ: Wiley.
- Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwanto, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act (Research report)*. Washington, DC: Urban Institute.
- Carnegie Council on Adolescent Literacy (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York.
- Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*, 278–293.
- Chall, J. (1983). *Stages of reading development*. New York, NY: McGraw Hill.
- Cosentino de Cohen, C., Deterding, N., & Chu Clewell, B. (2005). *Who's left behind? Immigrant children in high and low LEP schools (Policy Report)*. Washington, DC: Urban Institute.
- Crawford, J. (2004). *Educating English learners: Language diversity in the classroom* (5th ed.). Los Angeles, CA: Bilingual Education Services.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimal age question and some other matters. *Working Papers on Bilingualism, 19*, 121–129.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934–945.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language*. Baltimore, MD: Brookes.
- Duncan, S. E., & DeAvila, E. A. (1998). *Pre-language assessment scales [PreLAS]*. Monterey, CA: CTB/McGraw-Hill.
- Duncan, S. E., & DeAvila, E. A. (2000). *PreLAS 2000 Technical Report, Forms C & D*. Monterey, CA: CTB/McGraw-Hill.
- Francis, D. J., Lesaux, N. K., & August, D. (2006). Language of instruction. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, New Jersey: Lawrence Erlbaum.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, growth curves analysis. *Journal of Educational Psychology, 88*, 3–17.
- Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives, 11*(36) Retrieved from <http://epaa.asu.edu/epaa/v11n36/>
- García, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly, 26*, 371–392.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd edition). New York, NY: Lawrence Erlbaum.
- Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (2006). *Education English language learners: A synthesis of research evidence*. Cambridge, UK: Cambridge University Press.
- Geva, E., & Genesee, F. (2006). First-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, New Jersey: Lawrence Erlbaum.
- Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*, 330–344.
- Gutierrez, K. D., Zepeda, M., & Castro, D. C. (2010). Advancing early literacy learning for all children: Implications of the NELP report for dual-language learners. *Educational Researcher, 39*, 334–339.
- Hutchinson, J. M., Whiteley, H. E., Smith, C. D., & Connors, L. (2003). The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading, 26*, 19–32.
- Jetton, T. L., & Alexander, P. E. (2004). Domains, teaching, and literacy. In T. Jetton (Ed.), *Adolescent literacy research and practice* (pp. 15–39). New York, NY: Guilford Press.
- Jongejan, W., Verhoeven, L., & Siegel, L. S. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology, 99*, 835–851.
- Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*, 851–868.
- Kieffer, M. J. (2010). English proficiency, socioeconomic status, and late-emerging reading difficulties. *Educational Researcher, 39*, 484–486.
- Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology, 31*, 475–483.
- Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal, 47*, 596–632.
- Lesaux, N. K., Koda, K., Siegel, L. S., & Shanahan, T. (2006). Development of literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 75–122). Mahwah, NJ: Lawrence Erlbaum.
- Lipka, O., & Siegel, L. S. (2007). The development of reading skills in children with English as a second language. *Scientific Studies of Reading, 11*, 105–131.
- Lonigan, C. J., & Shanahan, T. (2010). Developing early literacy skills: Things we know we know and things we know we don't know. *Educational Researcher, 39*, 340–346.
- MacSwan, J., & Rolstad, K. (2006). How language proficiency tests mislead us about ability: Implications for English language learner placement in special education. *Teachers College Record, 108*, 2304–2328.
- MacSwan, J., Rolstad, K., & Glass, G. V. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-LAS Español. *Bilingual Research Journal, 26*, 395–421.
- Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology, 102*, 701–711.
- Manis, F., Lindsey, K., & Bailey, A. (2004). Development of reading in grades K-2 in Spanish-speaking English-language learners. *Learning Disabilities Research and Practice, 19*(4), 214–224.
- Najarian, M., Pollack, J. M., & Sorongon, A. G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K)*. Psychometric report for the eighth grade (NCES 2009–002). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading & Writing: An Interdisciplinary Journal, 20*, 691–719.
- Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading, 12*, 351–371.
- National Academy of Education (2009). *Education policy briefing sheet: Literacy for learning*. National Academy of Education White Papers Project. Washington, DC: Authors.
- National Center for Educational Statistics (2009). *Nation's Report Card: Reading, 2009*. Washington, DC: Author.
- National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute of Education.
- Orellana, M. F., & D'warte, J. (2010). Recognizing different kinds of "head starts.". *Educational Researcher, 39*, 295–300.
- RAND Reading Study Group (2002). *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. Santa Monica, CA: RAND Corporation.
- Reese, L., Garnier, H., Gallimore, R., & Goldenberg, C. (2000). Longitudinal analysis of the antecedents of emergent Spanish literacy and middle-school English reading achievement of Spanish-speaking students. *American Educational Research Journal, 37*, 633–662.
- Schnatschneider, C., & Lonigan, C. J. (2010). Misunderstood statistical assumptions undermine criticism of the National Early Literacy Panel's report. *Educational Researcher, 39*, 347–351.
- Senechal, M., & LaFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*, 445–460.
- Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*, 247–284.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Snow, C. E., Griffin, P., & Burns, M. S. (Eds.). (2005). *Knowledge to support the teaching of reading*. San Francisco, CA: Jossey-Bass.
- Snow, C. E., Porche, M. V., Tabors, P. O., & Harris, S. R. (2007). *Is literacy enough? Pathways to academic success for adolescents*. Baltimore, MD: Paul H. Brookes.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*, 934–947.
- Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology, 46*, 413–429.

- Tourangeau, K., Lê, T., Nord, C., & Sorongon, A. G. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K)*. Eighth-grade methodology report (NCES 2009–003). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- U.S. Department of Education, National Center for Education Statistics (2002). *Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS-K)*, Psychometric Report for Kindergarten Through First Grade, NCES 2002–05, by D. A. Rock and J. M. Pollack, Educational Testing Service, Elvira Germino Hausken, project officer. Washington, DC.
- Verhoeven, L. T. (1994). Transfer in bilingual development: The linguistic interdependence hypothesis revisited. *Language Learning*, 44, 381–415.
- Wiig, E. H., Semel, E. M., & Secord, W. (1992). *CELF-Preschool: Clinical evaluation of language fundamentals—Preschool*. San Antonio, Texas: Psychological Corporation, Harcourt Brace Jovanovich.